Special Issue

Article

http://dx.doi.org/10.21577/0103-5053.20180234

J. Braz. Chem. Soc., Vol. 30, No. 3, 499-508, 2019
Printed in Brazil - ©2019 Sociedade Brasileira de Química

Heterologous Expression of a Putative ClpC Chaperone Gene Leads to Induction of a Host Metabolite

Jana Braesel [a] ***and Alessandra S. Eustáquio*** *,[a]

[a]*Department of Medicinal Chemistry and Pharmacognosy and Center for Biomolecular Sciences, College of Pharmacy, University of Illinois at Chicago, 900 South Ashland Ave., 60607 Chicago-IL, USA*

Genome mining provides exciting opportunities for the discovery of natural products. However, in contrast to traditional bioassay-guided approaches, challenges of genome mining include poor or no expression of biosynthetic gene clusters (BGCs). Additionally, given that thousands of BGCs are now available through extensive genome sequencing, how does one select BGCs for discovery? Synthetic biology techniques can be used for BGC refactoring and activation, whereas resistance-gene-directed genome mining is a promising approach to discover bioactive natural products. Here we report the selection of a BGC by applying a resistance-gene-directed approach, cloning of the silent BGC from *Micromonospora* sp. B006, promoter exchange, and heterologous expression in *Streptomyces coelicolor* M1152. While we have yet to identify the encoded compound, we unexpectedly observed induction of a host metabolite, which we hypothesize is due to the presence of a caseinolytic protease C (ClpC) chaperone gene in the BGC, suggesting that ClpC chaperones may be used for BGC activation.

Keywords: Clp protease, Clp ATPase, Clp chaperone, genome mining, gene activation, natural product, antibiotic

Introduction

Natural products are an exceptional source of drug leads. For instance, the vast majority of antibiotics and anticancer agents in the clinic today are natural products or derivatives thereof. Given the rise of drug resistance, there is an urgent need to discover new drugs, and natural products have proven to be a reliable source.[1]

In spite of already being prolific sources, whole genome sequencing has shown that the biosynthetic potential of microorganisms has been underestimated based on the small number of compounds that are produced by each organism under routine laboratory growth conditions.[2,3] In bacterial genomes, genes for the biosynthesis of natural products are usually co-localized in so-called biosynthetic gene clusters (BGCs). With the fast developments in genome sequencing, public databases now contain thousands of orphan BGCs, that is, BGCs for which the corresponding compound is unknown. Thus, genome mining is a promising approach for the discovery of new natural products.[4] However, in contrast to traditional bioassay-guided discovery, challenges of genome mining include the fact that some BGCs are poorly or not at all expressed under a particular laboratory growth condition (so-called "silent" BGCs), preventing compound identification and isolation. Additionally, given the wealth of BGCs available, how does one select BGCs for discovery?

Genetic engineering and synthetic biology approaches can help with BGC activation. For example, promoter replacements have been shown to lead to the activation of silent BGCs.[5] Heterologous expression has also been used as an alternative approach to the genetic engineering of native producers.[6] Moreover, resistance-gene-directed genome mining is a promising approach to discover bioactive natural products.[7-9] Resistance-gene-directed genome mining relies on the observation that most antibiotic BGCs contain not only genes encoding enzymes to catalyze antibiotic biosynthesis, but also regulatory genes and resistance genes that prevent self-killing. Self-resistance mechanisms include export pumps, antibiotic-modifying enzymes, and target modification.[10] In the case of self-resistance genes that encode target modification, the resistance gene already gives an indication of the antibiotic target, the reason this approach has also been termed target-directed genome mining.[7] For instance, the BGC for the gyrase inhibitor novobiocin contains a gene

*e-mail: ase@uic.edu

encoding a gyrase subunit that was shown to be resistant to novobiocin.[11] Similarly, the BGC for the proteasome inhibitor salinosporamide A contains a gene encoding a salinosporamide-resistant proteasome subunit.[12] The resistance gene present in the BGC is in addition to a housekeeping copy present elsewhere in the genome. Duplication and differential regulation are presumably important because the resistant copy may be less active than the housekeeping one; in other words, resistance may have evolved at the expense of a loss in enzyme activity. Thus, the target-directed genome mining approach looks for duplication of housekeeping, essential genes, and the association of one of the duplicated copies with a natural product BGC.[7,8]

In the examples mentioned above the compounds and their modes of action were known before the BGC was identified. With the advancements in genome sequencing, bioinformatics, and synthetic biology, the scientific community is now in a position to potentially identify BGCs that encode compounds for a specific target of interest before the compound is even isolated, as has recently been demonstrated.[7,9]

One of the targets we are interested in is the caseinolytic protease (Clp) system. Clp proteases have been identified as promising antibacterial targets.[13] They play a central role in a great number of processes ranging from general maintenance of protein quality to tight control of key regulatory elements such as transcription factors with impact on stress survival, virulence, and antibiotic resistance.[14] The Clp complex is composed of a proteolytic core (assembled from 14 caseinolytic protease P (ClpP) subunits) and various adenosine triphosphatase (ATPase) orthologs (e.g., caseinolytic proteases X, C or A (ClpX, ClpC or ClpA, respectively)).[15] In association with the proteolytic core, the Clp ATPases are responsible for recognition, unfolding, and translocation of the substrate into the ClpP degradation chamber.[16] Independently of ClpP, they support the refolding process of aberrant proteins. In *Mycobacterium tuberculosis* and other related Actinobacteria, ClpP, ClpX, and ClpC are essential enzymes.[14] An interesting finding in recent years was the discovery of several new types of antibiotics, such as cyclomarin A, lassomycin and ecumicin, which kill bacteria by interfering with the function of Clp protease subunit ClpC1.[17-21] They show potent activity against multidrug-resistant *M. tuberculosis* strains, as well as drug-sensitive strains.[17,18,20,22]

Here we report the selection of a putative ClpC-containing BGC from *Micromonospora* sp. strain B006, its cloning, promoter exchange, and heterologous expression in *Streptomyces coelicolor* M1152, which unexpectedly led to induction of a host metabolite.

Experimental

General experimental procedures

All chemicals were acquired from Sigma-Aldrich, Alfa Aesar, MP Biomedicals, VWR, and Fisher Scientific. Solvents were of high performance liquid chromatography (HPLC) grade or higher. Restriction enzymes were purchased from New England Biolabs. Oligonucleotide primers were synthesized by Sigma-Aldrich. Molecular biology procedures were carried out according to the manufacturers' instructions (Ambion, Invitrogen, New England Biolabs, Thermo Fisher Scientific, Sigma-Aldrich, Qiagen, Zymo Research) or as described below.

Cultivation conditions

Micromonospora sp. B006 was routinely cultivated at 30 °C on International Streptomyces Project 2 (ISP2) medium (0.4% yeast extract, 1% malt extract, 0.4% dextrose, pH 7.3, and 2% agar for solid medium). To isolate genomic deoxyribonucleic acid (DNA), strain B006 was cultivated in a 1:1 mixture of tryptic soy broth (TSB) medium (3% tryptic soy broth) and yeast extract-malt extract (YEME) medium (0.3% yeast extract, 0.5% peptone, 0.3% malt extract, 1% glucose, 34% sucrose, 5 mM $MgCl_2$) for two days at 30 °C and 200 rpm.

Streptomyces coelicolor M1152 was maintained on mannitol soya flour (MS) solid medium (20 g soy flour, 20 g mannitol, and 20 g agar in 1 L tap water).[23] To isolate genomic DNA, *S. coelicolor* M1152 wild-type strain and exconjugants were cultivated in TSB liquid medium for 3 days at 30 °C and 200 rpm. To obtain spores for conjugation, *S. coelicolor* M1152 was cultivated at 30 °C on MS solid medium for two to three weeks.

Escherichia coli strains were cultivated in lysogeny broth (LB) medium supplemented with the appropriate antibiotics. The following antibiotics were used as selection markers: kanamycin (50 μg mL^{-1}), chloramphenicol (25 μg mL^{-1}), and nalidixic acid (25 μg mL^{-1}).

Construction of capture vector pJB005EL for transformation-associated recombination (TAR) cloning in yeast

The 119 base pair (bp) double-stranded DNA (*ds*DNA) fragments were designed as shown in Figure 1 (5'-CTCGAGGCAGGATCACCGTGCAGGAGGCGACGGCTCGCATCGCTGACCTGGTCACCGGATCCCCAACCTTGGTCAGTCATGACCCACGAGCGTTAGAAGCAGGACTGTCTCGCCATATG-3'; restriction sites are underlined; C: cytosine, T: thymine, G: guanine and

A: adenine) and synthesized by GenScript. The two restriction sites *Xho*I and *Nde*I were used to clone the synthetic DNA fragment into the same sites of the vector pCAP03-*acc(3)IV*[7] yielding the capture vector pJB005EL. Prior to direct TAR cloning, the capture vector was digested with *Bam*HI.

Cloning of the non-ribosomal peptide synthetase (NRPS)-polyketide synthase (PKS) gene cluster by direct capture from genomic DNA using TAR

High molecular weight genomic DNA was isolated from *Micromonospora* sp. B006 using the Blood & Cell Culture DNA Midi Kit (Qiagen). Direct TAR cloning of the NRPS-PKS gene cluster from genomic DNA was carried out using a previously reported protocol with minor modifications.[7] Solutions used were prepared according to Kouprina and Larionov.[24] Spheroplast cells were prepared using Zymolyase (100T equivalent; Zymo Research) at a final concentration of 1 mg mL^{-1}. The Zymolyase reaction was incubated for 50 min at 30 °C and 70 rpm. A total of 3 µg of genomic DNA and 1 µg of the linearized capture vector were used for yeast transformation. The transformed spheroplasts were mixed with 9 mL of synthetic tryptophan drop-out (SD-Trp) top agar (containing 1.5% agar) at 55 °C and overlaid on SD-Trp containing 5-fluoroorotic acid (FOA) agar (containing 2% agar). The plates were incubated for 6 days at 30 °C. Transformants were picked and transferred onto new SD-Trp-FOA agar plates and incubated for 3 days at 30 °C. Genomic DNA extraction from yeast was carried out using single-tube lithium acetate-sodium dodecyl sulfate (SDS) lysis.[25] A single colony of each transformant was suspended in 100 µL of a solution containing 200 mM lithium acetate and 1% SDS. The reactions were vortexed and then incubated for 5 min at 70 °C. Subsequently, 300 µL of absolute ethanol (molecular biology grade) was added for DNA precipitation. Samples were mixed by inverting, incubated for 15 min at –20 °C and then centrifuged for 10 min at 15,000 rpm. The pellets were then washed with 70% ethanol. After the pellets were dried, they were resuspended in 100 µL sterile distilled water. The transformants were screened by polymerase chain reaction (PCR) using primer pairs oJB7/oJB8 and oJB9/oJB10, respectively (Table 1). The 20 µL reactions consisted of 0.2 mM of each deoxynucleoside triphosphate (dNTP), 3% dimethylsulfoxide (DMSO), 0.5 µM (each) primer, and 0.02 U $µL^{-1}$ Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) in the buffer supplied with the enzyme and using the following thermal cycling parameters: 30 s at 98 °C; 35 cycles of 98 °C for 10 s, 62.8/65 °C for 20 s, and 72 °C for 30 s; and a terminal extension for 5 min at 72 °C. Plasmid pJB005EL was used as negative control. PCR positive clones were then grown in 4 mL SD-Trp liquid medium for 35 h at 30 °C and 220 rpm. Cells were lysed using Zymolyase (100T equivalent) at 30 °C for 2 h. Plasmids were extracted following the protocol of Tang *et al.*[7] and then transferred into *E. coli* DH5α by electroporation. The plasmids were purified from kanamycin-resistant *E. coli* clones using ZR BAC DNA Miniprep kit (Zymo Research) and then confirmed by restriction digest with *Bam*HI. The vector containing the NRPS-PKS hybrid gene cluster was designated as pJB007EL.

Promoter replacement

To ensure transcription in *S. coelicolor* M1152, the native promoter was replaced with the constitutive promoter

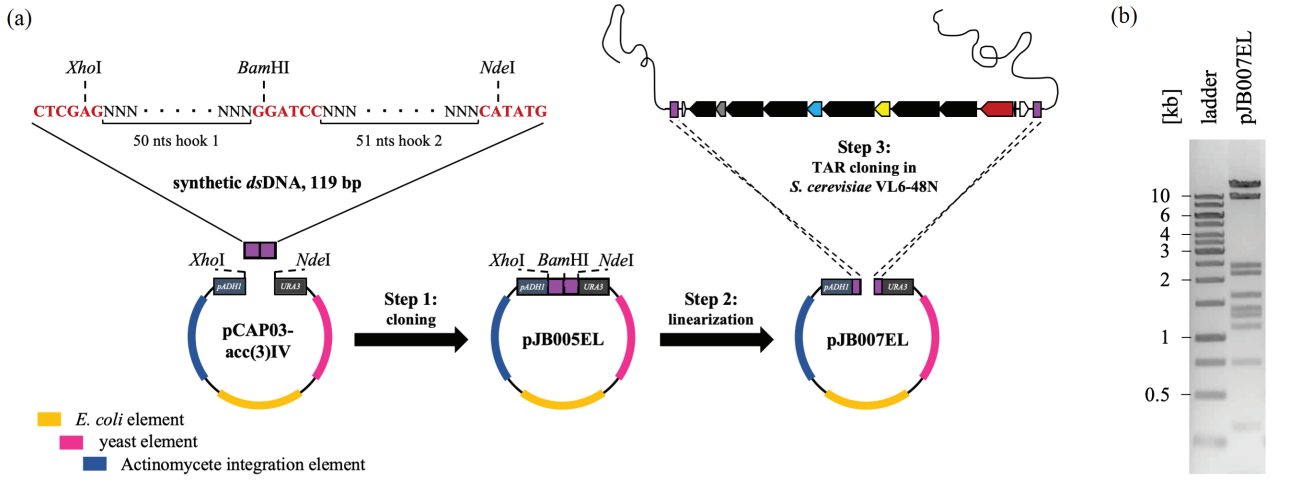

Figure 1. Cloning of the ClpC-containing BGC. (a) Schematic representation of cloning directly from genomic DNA using TAR in yeast; (b) confirmation of obtained clones by restriction digest with *Bam*HI. Expected fragments: 319, 740, 1,130, 1,290, 1,385, 1,630, 2,172, 2,397, 10,129, and 16,868 bp.

Table 1. Oligonucleotide primers used during this study

Primer[a]	Sequence (5' to 3')
P1_apra_f	ATG AGC TCA GCC AAT CGA C
P1_ORF_apra_f	CCG CCG CCT TTC GAT CAC GCC TTC GCC GCC CGA ACT CGG CAT GAG CTC AGC CAA TCG AC
P2_apra_r	GAT GTC ACG CTG AAA ATG C
P3_apra_ermE*_f	CGG CAT TTT CAG CGT GAC ATC GCT TGC AGG TCC AGG AAG
P4_apra_ermE*_f	ACC GCT GGA TCC TAC CAA C
P4_ORF_apra_ermE*_r	GGG AGG CGC GGG ACC CCG GTA CCC GCG GAC CCT GCG GTC TAC CGC TGG ATC CTA CCA AC
oJB7	CAC TCC ATC GTA TCC GTA TTC C
oJB8	GTT GAG ACT ACG CAT CGA CC
oJB9	AGC GTA CCG AGC AGC TTG C
oJB10	CTC TTC GCA ATG TCA ACA GTA CC
oJB71	GCA GGT CGT CGC GTT CAG
oJB72	CTC CAC GTC AGC CTG CCA
oJB76	AGC ACT TCG ATC GTG TCG
oJB77	GCT CCT TGG ACG ACT TGA
oJB78	CAC CAC CAC GTT CGC AAT C
oJB79	GGT TGG TTA CCG ACG ACG
oJB82	AGT GCG GTA GCA CCG CTC
oJB83	ACG CGG ATC TCT TCG CAC
oJB184	AGG CCC TTC GGG GTA CTC G
oJB185	GTG TTG CCG ACT TTC GTG ACG

[a]Overhangs are underlined. A: adenine; T: thymine; G: guanine; C: cytosine.

*ermE**[26,27] using λ-Red-mediated recombination[28-30] (Figure 2). Primer pair P1_apra_f/P2_apra_r (Table 1) was used to amplify the apramycin resistance gene from vector pIJ773,[31] while primers P3_apra_ermE*_f and P4_apra_ermE*_r were used to amplify the promoter *ermE** from vector pUWL201.[32] The 100 µL PCR reactions consisted of 0.2 mM of each dNTP, 3% DMSO, 0.25 µM of each primer, and 1 U μL^{-1} Phusion High-Fidelity DNA polymerase in HF reaction buffer supplied with the enzyme. Thermocycling parameters were: initial denaturation for 1 min at 98 °C; amplification: 30 cycles (98 °C for 10 s, 62 °C for 30 s, 72 °C for 30 s); and terminal hold for 5 min at 72 °C. The two obtained PCR fragments (936 bp *apr*^{R}, 217 bp P*ermE**) were purified using QIAquick PCR purification kit (Qiagen) and then digested with *Dpn*I for 2 h at 37 °C. Subsequently, the purified PCR fragments were ligated by overlap extension (OE) PCR using primers P1_ORF_apra_f and P4_ORF_apra_ermE*_r (Table 1). The 50-µL reaction consisted of the two purified DNA fragments in equimolar amounts (10 ng of *apr*^{R}, 4 ng of P*ermE**), 0.2 mM of each dNTP, 3% DMSO, and 1 U μL^{-1} Phusion High-Fidelity DNA polymerase in HF reaction buffer supplied with the enzyme. Thermocycling conditions were initial denaturation: 60 s at 98 °C; amplification: 3 cycles (98 °C for 10 s, 70 °C for 30 s, 72 °C for 90 s). Subsequently, 0.25 mM of each primer were added to the reaction. Thermocycling parameters were initial denaturation for 60 s at 98 °C; amplification: 30 cycles (98 °C for 10 s, 63.5 °C for 30 s, 72 °C for 90 s); and terminal hold for 5 min at 72 °C. Subsequently, the 1,233 bp PCR fragment was purified using the QIAquick gel extraction kit (Qiagen). The primers P1_ORF_apra_f and P4_ORF_apra_ermE*_r introduced two 40 bp overhangs for λ-Red recombination. In parallel, plasmid pJB007EL was transferred into *E. coli* BW25113/pIJ790 by electroporation. Since the plasmid pIJ790 is temperature sensitive, all incubation steps were performed at 30 °C. Positive clones were confirmed by restriction digest with *Bam*HI. *E. coli* BW25113/pIJ790/pJB007EL was grown in LB supplemented with kanamycin and chloramphenicol at 30 °C and 200 rpm until an optical density at 600 nm (OD_{600}) of 0.4 was reached. Then 10 mM L-arabinose was added to induce the λ-Red genes and the culture was incubated for an additional 30 min. The cells were washed twice with ice-cold 10% glycerol and resuspended in 1 mL 10% glycerol after centrifugation. Electro-competent cells (50 µL) were transformed with 100 ng purified OE-PCR product. Shocked cells were incubated in 1 mL LB at 37 °C for 1 h and plated on LB containing apramycin.

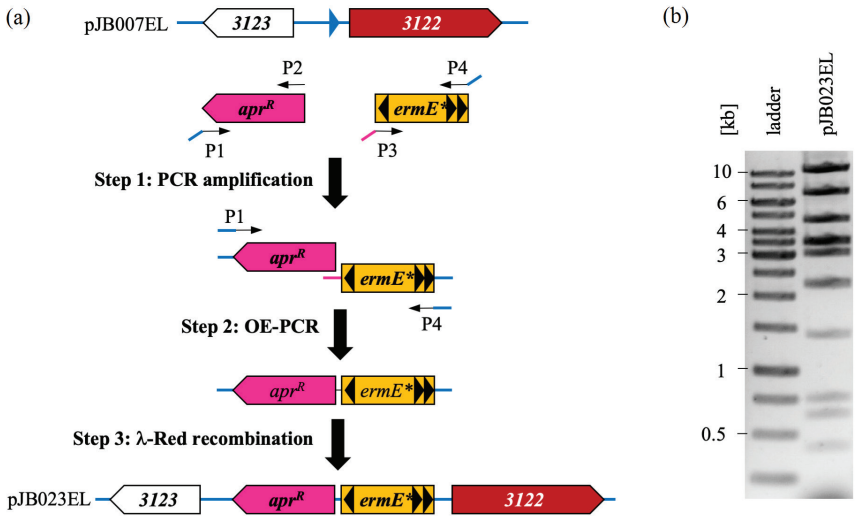

Figure 2. Promoter replacement. (a) Schematic representation of the promoter replacement design; (b) confirmation of obtained clones by restriction digest with *Sac*I. Expected fragments: 411, 627, 752, 1,419, 2,245, 3,070, 3,546, 3,574, 4,774, 7,263, and 11,307 bp. Triangles indicate promoter sequences.

Apramycin-resistant clones were grown for 6 h at 37 °C and 200 rpm. Subsequently, plasmids were isolated using ZR Plasmid Miniprep kit (Zymo Research) and confirmed by restriction digest with *Sac*I after passage into *E. coli* DH5α. The plasmid containing the NRPS-PKS hybrid gene cluster under control of the promoter *ermE** was designated as pJB023EL.

Heterologous expression of the NRPS-PKS gene cluster

The plasmid pJB023EL was transferred into *S. coelicolor* M1152 by conjugation from *E. coli* ET12567/pUZ8002 as described by Kieser *et al.*[23] Apramycin-resistant colonies were streaked on MS plates containing apramycin and nalidixic acid until pure cultures were obtained. Genomic DNA was isolated using GenElute™ Bacterial Genomic DNA kit (Sigma-Aldrich). Gene cluster integration was confirmed by PCR with primer pair oJB71/oJB72 (Table 1) in reactions containing 0.2 mM each dNTP, 3% DMSO, 0.2 µM (each) primer, and 1.25 U DreamTaq Green DNA polymerase (Thermo Fisher Scientific) in the buffer supplied with the enzyme, in a total volume of 20 µL. Thermocycling conditions were initial denaturation: 3 min at 95 °C; amplification: 35 cycles (95 °C for 30 s, 72 °C for 20 s, 72 °C for 60 s); terminal extension: 5 min at 72 °C. Genomic DNA isolated from the wild-type strain was used as negative control. Plasmid pJB023EL served as positive control.

Fermentation and metabolite analysis

For heterologous expression, 50 mL TSB liquid medium was inoculated with 400 µL of a frozen stock of *S. coelicolor* M1152 wild-type strain and exconjugants, respectively. Frozen stocks were prepared with mycelium from 3 days-old TSB liquid cultures by adding glycerol to 20% v/v final concentration followed by storage at –80 °C. The seed cultures were incubated for 3 days at 30 °C and 200 rpm. Subsequently, 5% v/v of these seed cultures were used to inoculate 50 mL of chemically defined medium (CDM) (30 g glucose, 6 g sodium citrate, 6 g L-proline, 2 g K_2HPO_4, 1.5 g $(NH_4)_2SO_4$, 5 g NaCl, 1 g $MgSO_4$, 0.4 g $CaCl_2$, 0.2 g $Fe_2SO_4 \cdot 7H_2O$, and 0.1 g $ZnSO_4 \cdot 7H_2O$ in 1 L distilled water, pH 7.2)[33] in a 250 mL Erlenmeyer flask. The production cultures were incubated at 30 °C and 200 rpm. After 7 days, 5% activated Diaion HP-20 resin (Alfa Aesar) was added to the production cultures. The HP-20 resin was activated prior to use by soaking in methanol and then rinsing thoroughly with distilled water. The production cultures were incubated for an additional 24 h at 30 °C and 200 rpm. The pH of the production fermentations was then adjusted to 4 with acetic acid. Subsequently, the cultures were harvested by centrifugation. The supernatant was decanted and the cell/resin pellet was extracted three times, each with 20 mL methanol. After removing the solvent under reduced pressure, the crude extracts were dissolved in methanol for HPLC and liquid chromatography (LC)-mass spectrometry (MS) analyses.

HPLC analysis was performed on an Agilent 1260 Infinity system equipped with a Kinetex® C18 column (150 × 4.6 mm, 5 µm particle size, 100 Å pore size, Phenomenex®). Solvent A was 0.1% v/v trifluoroacetic acid (TFA) in water, and solvent B was acetonitrile. The solvent gradient was: initial hold at 10% B for 1 min, linear gradient from 10 to 50% B within 34 min, followed by another linear gradient from 50 to 100% B within 15 min, and held for 4 min, at a flow rate of 1.0 mL min^{-1}. The detection

wavelength range was 200-600 nm; chromatograms were extracted at λ = 210, 254, 280, 310, 410, and 580 nm.

MS and MS/MS analyses were performed on a Finnigan LCQ Advantage MAX mass spectrometer system (Thermo Electron Corporation) in positive mode and a Hewlett Packard series 1050 HPLC, equipped with a Kinetex® 5 µm C18 100 Å-column (150 × 4.6 nm, 5 µm particle size, Phenomenex®), at a flow rate of 1 mL min^{-1}. Solvent A was 0.02% v/v formic acid in water, and solvent B was 0.02% v/v formic acid in acetonitrile. The gradient was: initial hold at 20% B for 1 min, linear gradient from 20 to 50% B within 20 min, followed by another linear gradient from 50 to 100% B within 1 min, and held for 3 min. The detection wavelength range was 410 nm. The detection mass range was 200 to 2,000 Da (positive mode).

Transcript analysis

S. coelicolor M1152 wild-type strain and exconjugant 2 were cultivated in 50 mL CDM liquid medium for 7 days at 30 °C and 200 rpm. Two-mL samples were taken on days 3, 5, and 7. The cells were collected by centrifugation and stored at –20 °C. Total ribonucleic acid (RNA) was isolated using the RiboPure™-Bacteria kit (Ambion). Cell lysis was obtained by homogenization using a Mini-Beadbeater™ (Biospec products) (5 cycles of 30 s). DNaseI treatment was carried out for 4 h using TURBO DNA-free™ kit (Ambion). The absence of genomic DNA in the samples was confirmed by PCR using 16S ribosomal RNA (rRNA) primers oJB184 and oJB185 (Table 1). The 20-µL reactions consisted of 0.2 mM of each dNTP, 3% DMSO, 0.25 µM of each primer, and 1.25 U DreamTaq DNA polymerase (Thermo Fisher Scientific) in reaction buffer supplied with the enzyme. Thermocycling parameters were: initial denaturation for 2 min at 95 °C; amplification: 30 cycles (95 °C for 30 s, 60 °C for 30 s, 72 °C for 60 s); and terminal hold for 5 min at 72 °C. Afterwards, complementary DNA (cDNA) synthesis was carried out with 250 ng RNA by using the SuperScript® IV First-Strand cDNA synthesis system (Invitrogen). A portion of the first-strand reaction (1 µL) was used as template in subsequent PCRs. For the conditions for the amplification of *MicB006_3120* (primer pair oJB71/oJB72), see above. The 20-µL reactions to amplify the cDNAs of *MicB006_3113* (primer pair oJB78/oJB79), *MicB006_3116* (primer pair oJB82/oJB83) and *MicB006_3122* (primer pair oJB76/oJB77) consisted of 0.2 mM of each dNTP, 1× Q5 High GC enhancer (New England Biolabs), 0.5 µM of each primer, and 0.02 U $µL^{-1}$ Q5 High-Fidelity DNA polymerase (New England Biolabs) in Q5 reaction buffer supplied with the enzyme. The following thermal cycling conditions were used: 30 s at 98 °C; 30 cycles of 98 °C for 10 s, 67.7/68.9/65 °C for 10 s, and 72 °C for 30 s; and a terminal hold for 2 min at 72 °C.

Results and Discussion

Resistance-gene-directed gene cluster selection

We have recently sequenced the complete genome of *Micromonospora* sp. strain B006 and shown that the 7-Mb genome encodes 16 BGCs, eleven of which remain orphan.[34] We were particularly intrigued by an NRPS-type I PKS gene cluster (cluster 11, genes *MicB006_3113* to *MicB006_3122*). Besides six genes encoding NRPSs and PKSs, the BGC also contains a gene encoding a putative ClpC ATPase, indicating that the gene cluster may encode a ClpC-targeting antibiotic (Figure 3). A basic local alignment search tool (BLAST)[35] search and subsequent phylogenetic analysis revealed nine genes encoding putative Clp subunits in the genome of strain B006, i.e., three ClpP proteases (*MicB006_0681*, *MicB006_3937*, and *MicB006_3938*), one ClpX subunit (*MicB006_3936*), one ClpB subunit (*MicB006_0157*), as well as four potential ClpC ATPases (*MicB006_1895*, *MicB006_4945*, *MicB006_3122*, and *MicB006_6016*). Thus, this BGC meets the requirements of target-directed genome mining, that is, it contains an extra copy of a housekeeping gene in association with natural product biosynthetic genes.

Cloning and promoter replacement

The 25.7-kb BGC (genes *MicB006_3113* to *MicB006_3122*) appeared transcriptionally silent under our culture conditions.[34] In order to activate its expression and to obtain the encoded compound, we cloned the entire BGC directly from the genomic DNA of *Micromonospora* sp. B006 by TAR cloning in yeast using the vector pCAP03-*acc(3)IV*[7] to yield pJB007EL (Figure 1a). The obtained plasmid was confirmed by restriction digest (Figure 1b).

To ensure transcription in the *S. coelicolor* heterologous host, we then replaced the native promoter with the *ermE**[26,27] constitutive promoter using λ-Red-mediated recombination[28-30] to generate pJB023EL (Figure 2a). Obtained clones were confirmed by restriction digest (Figure 2b).

Heterologous expression of the ClpC-containing BGC leads to induction of a host metabolite

S. coelicolor M1152 has been optimized by Gomez-Escribano and Bibb[36] for heterologous expression

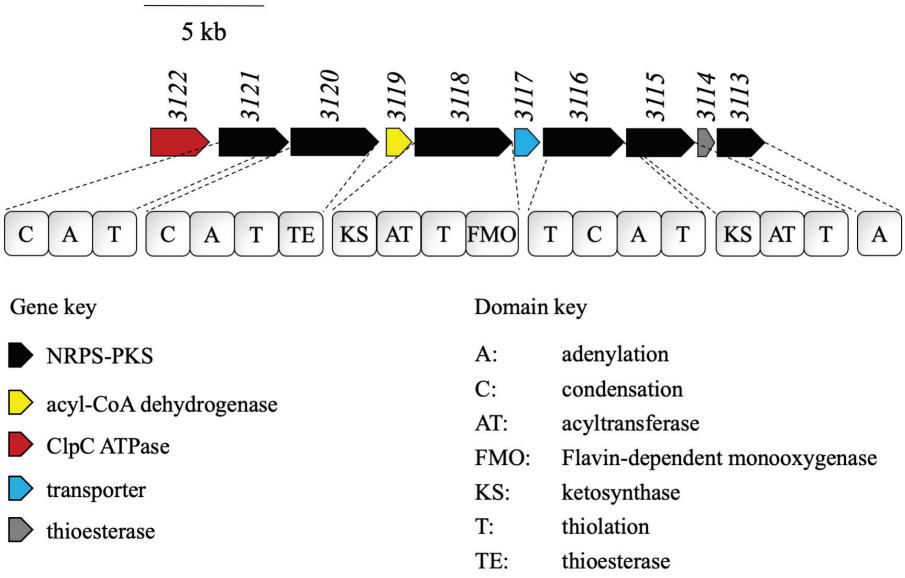

Figure 3. ClpC-containing biosynthetic gene cluster from *Micromonospora* sp. B006. Genes (arrows) are color-coded according to predicted function.

of natural product BGCs and was therefore chosen as host for our experiments. Plasmid pJB023EL was transferred into *S. coelicolor* M1152 by conjugation from *E. coli* ET12567/pUZ8002.[23] Mutants were confirmed by PCR. We then compared the metabolite profile of parent and mutant strains grown in CDM. A distinct peak at 30.8 min was observed in the mutants at a wavelength of 410 nm (Figure 4a). Upon closer inspection, this peak can be detected in the parent strain as well, although only as a trace compound.

The obtained UV-Vis spectrum of the peak with absorbance maxima at 334, 406, 536, and 572 nm showed high similarity to the UV-Vis spectrum of Zn coproporphyrin III available in our UV library (Figure 4a). Porphyrins are a class of photoactive molecules with a broad range of bioactivity.[37] They are widely used in photodynamic therapies.[38] The ability to produce porphyrins was reported for several *Streptomyces* spp.[37,39,40] Mass spectrometry analysis using liquid chromatography quadrupole (LCQ) showed an *m/z* of 407 for the induced compound, which appears doubly charged (Figure 4b). The isotope pattern indicates that no zinc is bound. The observed mass does not match Zn coproporphyrin III (calcd. for $C_{36}H_{36}N_4O_8Zn$: 716.182) or coproporphyrin III (calcd. for $C_{36}H_{38}N_4O_8$: 654.269). A search in the Global Natural Products Social Molecular Networking (GNPS)[41] led to no hits. A manual comparison of the obtained *m/z* with those of known porphyrins also led to no matches. We hypothesize that the induced compound may be an analog of known porphyrins. However, its structure remains to be determined.

Transcriptional analysis

Since we were unable to obtain the product of the NRPS-PKS gene cluster, we next analyzed gene transcription. Transcript analysis of four genes of the hybrid NRPS-PKS gene cluster, *MicB006_3113*, *MicB006_3116*, *MicB006_3120*, and *MicB006_3122*, revealed that only the gene encoding the putative ClpC ATPase was transcribed,

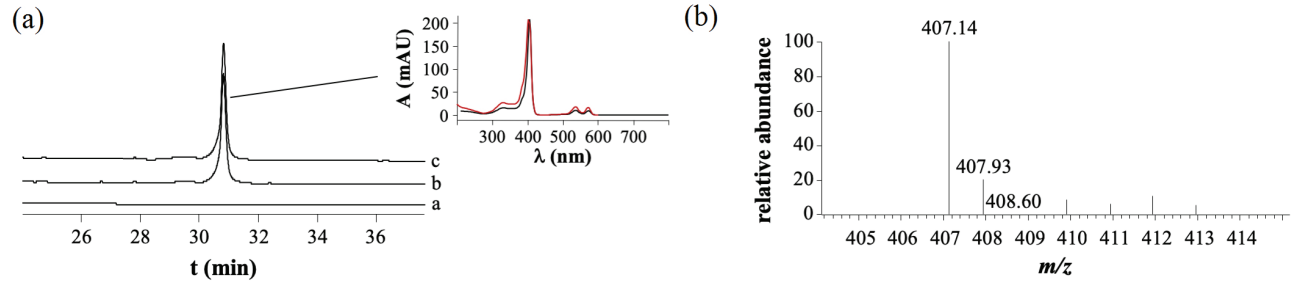

Figure 4. Heterologous expression. (a) HPLC analysis of culture extracts from parent strain *S. coelicolor* M1152 and from mutant strains (mt 2 and mt 7). The UV-Vis spectrum of the induced compound (black) was compared to Zn coproporphyrin III (red); (b) mass spectrometry analysis. Maximum intensity: 1.25×10^7. The spectrum shown is from mt 2. The same *m/z* is found in the parent strain as well.

which indicates that only this gene is functional in our heterologous system (Figure 5).

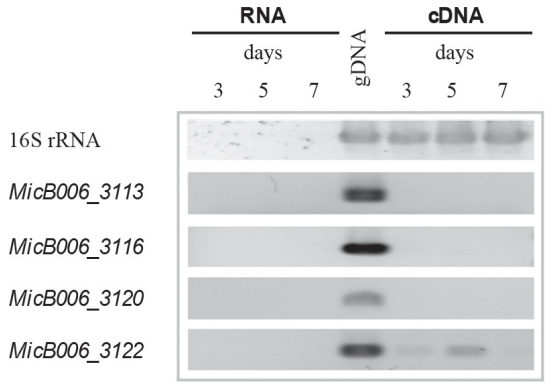

Figure 5. Transcriptional analysis. Reverse transcription PCR analysis of *S. coelicolor* containing pJB023EL grown in CDM medium.

Recent data support the idea that Clp ATPases could be involved in the activation of transcription.[42] In addition, ClpA, a ClpC homolog in gram-negative bacteria, performs an ATP-dependent chaperone function independent of ClpP.[43] It has been shown to remodel bacteriophage P1 RepA dimers into monomers, thereby activating the latent specific DNA binding activity of RepA.[43] In *Bacillus subtilis*, ClpC is involved in the switch that controls the activity of a competence-specific transcription factor (ComK).[44] ComK activates transcription of itself and other genes required for competence. When ClpC and MecA, which confers specific molecular recognition, bind to ComK in the presence of ATP, it results in the inhibition of ComK DNA binding. ComS, a small protein required for ComK biosynthesis, can resolve this ternary complex by interacting with MecA.[44] MecA and ClpC also positively regulate autolysin gene expression in a similar way in which they affect competence.[45] Here the effect is mediated through the sigma-D factor (SigD) protein instead of the ComK protein.[45]

Thus, our transcriptional analysis results combined with literature data suggest that expression of the ClpC gene is responsible for the induction of the host metabolite.

Conclusions

Freshwater Actinobacteria remain underexplored as a source of natural products. Resistance-gene-directed genome mining is a promising approach to discover bioactive compounds.[7,9] In aiming to apply a resistance-gene-directed approach for antibiotic discovery from freshwater Actinobacteria, we selected an NRPS-PKS BGC from a Lake Michigan *Micromonospora* sp. that contained a ClpC gene as a putative resistance gene. ClpC is an emerging anti-tuberculosis target. The *clpC1* gene of *M. tuberculosis* is predicted to be necessary for the survival of the organism.[46] The recently discovered antibiotics cyclomarin A, lassomycin, and ecumicin interact with ClpC1[17-20] and show potent activity against multidrug-resistant *M. tuberculosis* strains.[17,18,20,22]

Although heterologous expression of the ClpC-containing BGC in *S. coelicolor* M1152 did not lead to the production of the encoded compound, we unexpectedly observed induction of a host metabolite. This is interesting and intriguing to us because firstly, we are not aware of reported instances of host metabolite induction when attempting BGC heterologous expression. Secondly, since a significant proportion of BGCs identified in microbial genomes are silent, methods to activate expression and allow compound identification are of interest. Our results suggest that ClpC chaperone expression may lead to gene activation. We are currently working on elucidating the structure of the induced host metabolite, on investigating the mechanism of induction, and on engineering the BGC further to try and obtain the encoded compound.

Acknowledgments

We thank Jessica Cleary and Laura Sanchez for help with MS analysis. Financial support for this work was provided by the National Center for Advancing Translational Sciences, National Institutes of Health (NIH), under grant KL2TR002002 (to A. S. E.) and by startup funds from the Department of Medicinal Chemistry and Pharmacognosy and the Center for Biomolecular Sciences, University of Illinois at Chicago (to A. S. E.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

1. Newman, D. J.; Cragg, G. M.; *J. Nat. Prod.* **2016**, *79*, 629.
2. Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A.; *Nature* **2002**, *417*, 141.
3. Udwary, D. W.; Zeigler, L.; Asolkar, R. N.; Singan, V.; Lapidus, A.; Fenical, W.; Jensen, P. R.; Moore, B. S.; *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10376.

4. Ziemert, N.; Alanjary, M.; Weber, T.; *Nat. Prod. Rep.* **2016**, *33*, 988.
5. Rutledge, P. J.; Challis, G. L.; *Nat. Rev. Microbiol.* **2015**, *13*, 509.
6. Baltz, R. H.; *J. Ind. Microbiol. Biotechnol.* **2010**, *37*, 759.
7. Tang, X.; Li, J.; Millan-Aguinaga, N.; Zhang, J. J.; O'Neill, E. C.; Ugalde, J. A.; Jensen, P. R.; Mantovani, S. M.; Moore, B. S.; *ACS Chem. Biol.* **2015**, *10*, 2841.
8. Alanjary, M.; Kronmiller, B.; Adamek, M.; Blin, K.; Weber, T.; Huson, D.; Philmus, B.; Ziemert, N.; *Nucleic Acids Res.* **2017**, *45*, W42.
9. Yan, Y.; Liu, Q.; Zang, X.; Yuan, S.; Bat-Erdene, U.; Nguyen, C.; Gan, J.; Zhou, J.; Jacobsen, S. E.; Tang, Y.; *Nature* **2018**, *559*, 415.
10. Munita, J. M.; Arias, C. A.; *Microbiol. Spectrum* **2016**, *4*, VMBF-0016-2015.
11. Steffensky, M.; Mühlenweg, A.; Wang, Z.-X.; Li, S.-M.; Heide, L.; *Antimicrob. Agents Chemother.* **2000**, *44*, 1214.
12. Kale, A. J.; McGlinchey, R. P.; Lechner, A.; Moore, B. S.; *ACS Chem. Biol.* **2011**, *6*, 1257.
13. Malik, I. T.; Brotz-Oesterhelt, H.; *Nat. Prod. Rep.* **2017**, *34*, 815.
14. Olivares, A. O.; Baker, T. A.; Sauer, R. T.; *Nat. Rev. Microbiol.* **2016**, *14*, 33.
15. Brotz-Oesterhelt, H.; Sass, P.; *Int. J. Med. Microbiol.* **2014**, *304*, 23.
16. Zolkiewski, M.; *Mol. Microbiol.* **2006**, *61*, 1094.
17. Schmitt, E. K.; Riwanto, M.; Sambandamurthy, V.; Roggo, S.; Miault, C.; Zwingelstein, C.; Krastel, P.; Noble, C.; Beer, D.; Rao, S. P.; Au, M.; Niyomrattanakit, P.; Lim, V.; Zheng, J.; Jeffery, D.; Pethe, K.; Camacho, L. R.; *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 5889.
18. Gavrish, E.; Sit, C. S.; Cao, S.; Kandror, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K.; *Chem. Biol.* **2014**, *21*, 509.
19. Bürstner, N.; Roggo, S.; Ostermann, N.; Blank, J.; Delmas, C.; Freuler, F.; Gerhartz, B.; Hinniger, A.; Hoepfner, D.; Liechty, B.; Mihalic, M.; Murphy, J.; Pistorius, D.; Rottmann, M.; Thomas, J. R.; Schirle, M.; Schmitt, E. K.; *ChemBioChem* **2015**, *16*, 2433.
20. Gao, W.; Kim, J. Y.; Anderson, J. R.; Akopian, T.; Hong, S.; Jin, Y. Y.; Kandror, O.; Kim, J. W.; Lee, I. A.; Lee, S. Y.; McAlpine, J. B.; Mulugeta, S.; Sunoqrot, S.; Wang, Y.; Yang, S. H.; Yoon, T. M.; Goldberg, A. L.; Pauli, G. F.; Suh, J. W.; Franzblau, S. G.; Cho, S.; *Antimicrob. Agents Chemother.* **2015**, *59*, 880.
21. Culp, E.; Wright, G. D.; *J. Antibiot.* **2017**, *70*, 366.
22. Schultz, A. W.; Oh, D. C.; Carney, J. R.; Williamson, R. T.; Udwary, D. W.; Jensen, P. R.; Gould, S. J.; Fenical, W.; Moore, B. S.; *J. Am. Chem. Soc.* **2008**, *130*, 4507.
23. Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A.; *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, 2000.
24. Kouprina, N.; Larionov, V.; *Nat. Protoc.* **2008**, *3*, 371.
25. Lõoke, M.; Kristjuhan, K.; Kristjuhan, A.; *BioTechniques* **2011**, *50*, 325.
26. Bibb, M. J.; Janssen, G. R.; Ward, J. M.; *Gene* **1985**, *38*, 215.
27. Myronovskyi, M.; Luzhetskyy, A.; *Nat. Prod. Rep.* **2016**, *33*, 1006.
28. Datsenko, K. A.; Wanner, B. L.; *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 6640.
29. Murphy, K. C.; Campellone, K. G.; Poteete, A. R.; *Gene* **2000**, *246*, 321.
30. Yu, D.; Ellis, H. M.; Lee, E. C.; Jenkins, N. A.; Copeland, N. G.; Court, D. L.; *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5978.
31. Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F.; *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 1541.
32. Steffensky, M.; Li, S. M.; Heide, L.; *J. Biol. Chem.* **2000**, *275*, 21754.
33. Kominek, L. A.; *Antimicrob. Agents Chemother.* **1972**, *1*, 123.
34. Braesel, J.; Crnkovic, C. M.; Kunstman, K. J.; Green, S. J.; Maienschein-Cline, M.; Orjala, J.; Murphy, B. T.; Eustaquio, A. S.; *J. Nat. Prod.* **2018**, *81*, 2057.
35. U.S. National Library of Medicine (NLM), National Center for Biotechnology Information (NCBI); *Basic Local Alignment Search Tool (BLAST)*; NLM/NCBI, USA, 2018.
36. Gomez-Escribano, J. P.; Bibb, M. J.; *Microb. Biotechnol.* **2011**, *4*, 207.
37. Martinez, A. F. C.; Almeida, L. G.; Moraes, L. A. B.; Consoli, F. L.; *BMC Microbiol.* **2017**, *17*, 143.
38. Ferreira, D. P.; Conceição, D. S.; Calhelha, R. C.; Sousa, T.; Socoteanu, R.; Ferreira, I.; Ferreira, L. F. V.; *Carbohydr. Polym.* **2016**, *151*, 160.
39. Musilek, V.; *Science* **1962**, *137*, 674.
40. Sato, K.; Shimizu, S.; Fukui, S.; *Agric. Biol. Chem.* **1968**, *32*, 1.
41. Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Crusemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderon, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C. C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C. C.; Yang, Y. L.; Humpf, H. U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya P, C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J.

A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodriguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P. M.; Phapale, P.; Nothias, L. F.; Alexandrov, T.; Litaudon, M.; Wolfender, J. L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D. T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Muller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutierrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N.; *Nat. Biotechnol.* **2016**, *34*, 828.

42. Burton, B. M.; Baker, T. A.; *Protein Sci.* **2005**, *14*, 1945.
43. Pak, M.; Wickner, S.; *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 4901.
44. Turgay, K.; Hamoen, L.; Venema, G.; Dubnau, D.; *Genes Dev.* **1997**, *11*, 119.
45. Rashid, M. H.; Tamakoshi, A.; Sekiguchi, J.; *J. Bacteriol.* **1996**, *178*, 4861.
46. Kar, N. P.; Sikriwal, D.; Rath, P.; Choudhary, R. K.; Batra, J. K.; *FEBS J.* **2008**, *275*, 6149.

Submitted: August 14, 2018
Published online: December 3, 2018